

Data Management In Cellular Networks Using Activity Mining

Dr. R. Ramachandran, Ph.D.

Principal

Dept. of CSE,SVCE

Chennai, India

rrama@svce.ac.in

V. Aishwaryalakshmi

PG Scholar

Dept. of CSE, SVCE

Chennai, India

neefavenkat@gmail.com

Abstract— In the recent technology advances, an increasing number of users are accessing various information systems via wireless communication. The majority users in a mobile environment are moving and accessing wireless services for the activities they are currently unavailable inside. We propose the idea of complex activity for characterizing the continuously changing complex behavior patterns of mobile users. For the purpose of data management, a complex activity is copy as a sequence of location movement, service requests, the coincidence of location and service, or the interleaving of all above. An activity may be composed of sub activities. Different activities may exhibit dependencies that affect user behaviors. We argue that the complex activity concept provides a more specific, rich, and detail description of user behavioral patterns which are very useful for data management in mobile environments. Correct exploration of user activities has the possible of providing much higher quality and personalized services to individual user at the right place on the right time. We, therefore, propose new methods for complex activity mining, incremental maintenance, online detection and proactive data management based on user activities. In particular, we develop pre-fetching and pushing techniques with cost sensitive control to make easy analytical data allocation. First round implementation and simulation results shows that the proposed framework and techniques can significantly increase local availability, conserve execution cost, reduce response time, and improve cache utilization.

Keywords-Activity mining, Prefetching,,pushing, proactive data management, Genetic Algorithm;

I. INTRODUCTION

Diverse mobile services and development in wireless networks have stimulated an enormous number of people to employ mobile devices such as cellular phones and portable laptops as their communications means. The most salient feature of wireless networks is mobility support, which enables mobile users to communicate with others regardless of location. Majority of users in a mobile environment do not travel at random. They navigate from place to place with specific purposes in mind. In many cases, the patterns of location movement and service invocation of mobile users targeting similar purposes show strong similarity. The common patterns of location movement may due to geographic relationships between

locations or service distribution. The regularity in service invocation may come from the dependencies between services or the proximity of service providers. It is potentially beneficial to discover such mobility and service patterns to make easy network and data management. We propose the idea of complex activity for characterizing the continuously changing complex behavior patterns of mobile users. A complex activity is a sequence of location movement, service requests, the coincidence of location and service, or the interleaving of all above. An activity may be composed of sub activities. Different activities may show dependencies that affect user behaviors. We argue that the activity concept provides a more specific, rich, and detail description of user behavioral patterns which are very useful for data management in mobile environments. Proper exploration of such activities enables data management system to predict the user's next move, intended service or both for providing much higher quality and personalized services to individual user at the right place on the right time. Such kind of advanced information services call for new methods of complex activity mining, incremental maintenance, online detection, and proactive data management based on user activities. We propose new pattern mining and patterns processing algorithms to the discovery and maintenance of complex activities in mobile environments. Furthermore, we develop pre-fetching, pushing, and handoff techniques with cost sensitive control to make easy predictive data allocation and personalized services. Preliminary implementation and simulation results demonstrate that the activity-based proactive data management strategies can significantly conserve execution cost, reduce response time, improve cache utilization, and increase local availability.

II. RELATED WORK

It has been known for quite sometime that user behavior patterns are important for effective mobile computing. The First stage along this line of research is the acquisition of user behavior. Among the various methods for learning user behavior, data mining is probably the most widely used technique. It is well suited for discovering hidden patterns from large volume of data such as transaction log. The mining of mobility patterns, in particular focus attention of some previously proposed work.

- User Mobility Profile (UMP) is a combination of historic records and predictive patterns of mobile terminals, which serves as fundamental information for mobility management and enhancement of Quality of Service (QoS) in wireless multimedia networks. UMP framework is developed for estimating service patterns and tracking mobile users, including descriptions of location, mobility, and service requirements. For each mobile user, the service requirement is estimated using a mean square error method. Moreover, a new mobility model is designed to characterize not only stochastic behaviors, but historical records and predictive future locations of mobile users as well. Therefore, it incorporates aggregate history and current system parameters to acquire UMP. In particular, an adaptive algorithm is designed to predict the future positions of mobile terminals in terms of location probabilities based on moving directions and residence time in a cell. The authors G. Resta and P. Santi, [1] have proposed their schemes about User Behavior model approach in mobile environments and I.F. Akyildiz and W. Wang, [2] describes "The Predictive User Mobility Profile Framework for Wireless environments.

- R.V. Mathivaruni and V.Vaidehi [3] proposes An activity based mobility prediction strategy using markv modeling for wireless networks which describes the foremost objective of a wireless network is to facilitate the communication of mobile users regardless of their point of attachment to the network. The system must discern the location of the mobile terminal, to afford flawless service to the mobile terminal. Mobility prediction is widely used to assist handoff management, resource reservation and service pre configuration. Prediction techniques that are currently used don't consider the motivation behind the movement of mobile nodes and incur huge overheads to manage and manipulate the information required to make predictions. This paper proposes an activity based mobility prediction technique that uses activity prediction and Markov modeling techniques to devise a prediction methodology that could make accurate predictions than existing techniques.

- On The effect of group mobility to data replication in Ad-hoc networks: In this paper, we address the problem of replica allocation in a mobile ad-hoc network by exploring group mobility. We first analyze the group mobility model and derive several theoretical results. In light of these results, we propose a replica allocation scheme to improve the data accessibility.

- A. Yamasaki, H. Yamaguchi, S. Kusumoto, and T. Higashino [4] design a Mobility-aware Data management (MoDA) scheme for mobile ad hoc networks (MANETs) composed by mobile nodes such as urban pedestrians and vehicles. By fully utilizing the knowledge about the trajectories of mobile nodes, MoDA determines how replicas of data are copied and transferred among mobile nodes to provide the required data accessibility. Experimental results have shown that MoDA could achieve the small number of data transfers among mobile nodes while keeping reasonable accessibility.

III. COMPLEX ACTIVITIES

A. Behavior pattern

We analyze the user behaviour patterns in Mobile environments and formally characterize the idea of complex activities. Users in a mobile environment travel from one place to another. The moving pattern of a user can naturally be modelled as sequences of locations visited by the user. Mobile users also invoke services one after another when travelling. Service patterns emerge if a sequence of services is repeatedly invoked by the same or different user. Some services are offered only at specific locations such as gas stations, restaurants, movie theatres, etc. In such cases, locations and service invocations always come in pairs. Based on the Analysis above, we can characterize the basic user behaviour Patterns into three categories:

Location-only patterns (L-type): Sequences of locations Those are repeatedly visited by mobile users.

Service-only patterns (S-type): Sequences of services Those are repeatedly invoked by mobile users.

Location-service patterns (LS-type): Sequences of Location-service pairs that are repeatedly visited and invoked by mobile users. Location-only patterns occur when users are engaged in Movement activities that are formed by geographical Constraints. Service-only patterns arise when users are embarked on invocation activities that are formed by service dependencies. Location-service patterns are commonly seen when there are strong associations between location and Service pairs. Sometimes, patterns are the result of personal habits. These patterns constitute the primitive activities in Mobile environments. Built on top of primitive activities, a complex activity can be a combination of any number of primitive or other complex activities as sub activities. Complex activities: A complex activity $A = \{a_1, a_2, \dots\}$ is a frequently occurring sequence of activities such That each a_i is either a primitive or a complex activity. Based on the recursive definition, a primitive activity is Considered as the simplest type of complex activity. Therefore, the term activity is used when there is no need to Distinguish between them.. We argue that the notion of Complex activities hold the key to effective data management in mobile environments for several reasons: Complex activities can be used to faithfully model User behaviour in mobile environments. By identifying the current activity of a mobile user, we can predict his/her next move on both location and service invocation with much higher accuracy than traditional way of relying on motion estimation and accumulated read/write statistics. Once the user behaviour can be predicted with high Accuracy, it becomes possible to provide proactive Services for the user. In other words, we can allocate the required data items and reserve the necessary Resources at the right place on the right time.

B. Service Architecture

Activity mining is the characterizing of continuously changing complex behavior patterns of mobile users. Our algorithm for activity mining is based on the popular Apriori algorithm to identify all primitive and complex activities of a user behavior. The activities identified by the mining algorithm

do not have any structure information except for sequences. The relationships between activities are not clear. They are not well suited for activity processing and data management. For this purpose we devise an efficient data structure named as activity tree for the representation and incremental maintenance of activities. After successful identification of activity patterns, we still need to recognize the current activity of a user. Once the activity of a user can be detected and prioritized, we are all set for presenting the proactive data management techniques namely

- (a) Proactive pushing
- (b) Predictive handoff
- (c) Precision pre fetching

Once the user behavior can be predicted with high accuracy, it becomes possible to provide Proactive services for the user. In other words, we can allocate the required data items and reserve the necessary resources at the right place on the right time. To facilitate such an activity-based proactive data management services, several issues must be properly addressed. These issues, in turn, pose significant challenges to information system designers and service providers.

Activity mining: We must be able to unearth the complex activities from large volumes of user behavior logs. This naturally demands effective algorithms for data analysis and activity mining.

Activity representation: Once the activities are identified, we need efficient structures for representing the activities to facilitate effective data management.

Activity maintenance: Due to highly transient nature of mobile environments, complex activities evolve dynamically over time. This calls for incremental maintenance techniques that can smoothly adapt to the changes.

Activity detection: After successful identification of activity patterns, we still need to recognize the current activity of a user. The online detection techniques must be very efficient. Otherwise, we may run the risk of late identification of an activity that a user is no longer engaged in

Proactive data management: To provide effective information services, proactive or predictive data management techniques are highly desirable with the knowledge of a user's possible next move or service invocation. The algorithm used is given below:

```
Function apriori_gen (Ak_1)
Ck=0;
for p 2 Ak_1 do
for q 2 Ak_1 do
if p:item1 =q:item1 ^ ... ^ p:itemk_2 = q:itemk_2 ^
p:itemk_1 !=q:itemk_1 then
Ck+=q:itemk_1;
end
end
for c 2 Ck do
for {k "1"}-subsets s of c do
if s ∈ Ak_1 then delete c from Ck;
end
end
return Ck;
Function contain Activity (t, c)
```

Input: $t = \{a_1; a_2; \dots; a_n\}$ is a transaction and $c = \{c_1; c_2; \dots; c_m\}$ is an activity candidate.

```
if c == 0; then return True;
if n < m then return False;
for (i = 1; i ≤ n - m, 1++) do
if action Activity Match{ ai; c1 } ^
contain Activity{ aip1; ...; an }, {c2; ...; cm} then
return True;
end
Return False;
Function action Activity Match (a, c)
Input: a is an action and c is a primitive activity.
If c:location == null ^ c:service == null then
Return a:location == c:location ^ a:service == c:service;
else if c:location == null then
return a:location == c:location;
else
return a:service == c:service;
```

C. Proactive Datamangement

Proactive data management can contain following steps

- Activity detection
- Activity weighting and ranking
- Data management strategies
- Activity maintenance

In the Activity Detection, User maintains the activity table that store the current move of the user from the activity tree index. In the Activity Weighting and Ranking, Ranking the possible move of the user by use of

1. Service distance(d)
2. Service data size(t)
3. Intensity(p)
4. Conformance ratio(γ)
5. Degree of sharing(w)
6. Minimal co occurrence(m)

In the data management strategies, we can find how to data transfer and maintenance, that is If cell provide L only-proactive pushing. Else If cell provide S only- Predictive handoff else If cell provide L & S-Precision pre fetching In the Activity Maintenance, they Update the behavior details in the behavior table. For an activity a and the set of sharing nodes W, we define the activity score of a as follows:

$$a.score = (d \times t) \times (\lambda(a_i) \times \alpha(\text{parent}(a_i)) / W)$$

Once the activity of a user can be detected and prioritized, we are all set for presenting the proactive data management techniques. In general, when a user is recognized to be engaged in an activity, we can predict with high probability the user's next move (including location and/or service invocation).

Proactive pushing—When the next move of a user is a service-only activity, we know what the user is likely to invoke but don't know where he/she is heading. Therefore, it is potentially beneficial to push the service data directly to the client cache such that it is immediately available when the service is actually invoked.

Predictive handoff—When the next step is a location-only activity, we know where the user is going with no knowledge of the service invocation.

Precision pre fetching—If the next activity is a location service pair, we know exactly the user's next move and what service the user is going to invoke.. With the strategies in mind, selecting proper data management operations becomes straightforward. For a user engaged in a complex activity, we check the activity tree to retrieve the next activity. If it is a service-only activity, we proactively push the service data toward the user. If it is a location-only activity, we contact the base station of the predicted location and send out the data used by the current service. Finally, if it is a location-service pair, we inform the base station of the next location to initiate the pre fetching of the predicted service data.

D Cost sensitive Operation

Even with careful activity identification and proper data management operation selection, it may not be always cost effective to apply the selected operations. The purpose of cost-sensitive operation control is to judiciously estimate the relative cost and benefit before applying the operations. They are only performed when the expected saving is higher than the management overhead. For proactive pushing, we push the predicted service data directly to the client cache. The pushing cost is, therefore, the service data transfer cost from the source to the client as follows: $S(fd+1)$ If the prediction is correct, the user's request can be satisfied immediately without further delay or cost. However, with probability $1-\alpha$, we may need the extra cost of waiting and transferring the data of the service actually invoked by the user when the prediction is wrong. This can be characterized as follows: $(1-\alpha)((D+R/Bfd+R/Bt)\lambda+R(fd+1))$

$$\frac{C}{D+T} < \frac{\alpha}{1-\alpha} \lambda.$$

Since α can be estimated by the conformance rate while C , D , and T can also be measured once the service data are identified, can be used for controlling the application of the proactive pushing strategy. Degree of sharing, a data management operation can benefit many users at the same time which provides a form of consideration on overall system. Thus system intergration.

IV. PROPOSED WORK

A. Adaptive Genetic Selection Metastatergies:

Met strategies for Adaptation that is quickly adjust the system when the situation changes, by making use of Adaptive generic algorithm to predict user next move further accurately. One more inclusion is activity prioritization that we can provide first priority to more important service than others in the complex user behavior. On the Adaptive Strategy Selection (AdapSS), in the Genetic Algorithms community, its objective is to automatically select between the available possibly ill-known mutation strategies, according to their performance on the current search/optimization process. To do so, there is the need for two components: the credit assignment, that defines how the impact of the strategies on the search should be assessed and transformed into a numerical reward; and the strategy (or operator) selection

mechanism that, based on the rewards received, select which strategy should be applied at the given moment of the search. Suppose we have $K > 1$ strategies in the pool $A = \{a_1, \dots, a_K\}$ and a probability vector $P(t) = \{p_1(t), \dots, p_K(t)\}$

$$(\forall t : p_{min} \leq p_i(t) \leq 1; \sum_{i=1}^K p_i(t) = 1).$$

In this work, the PM (Probability Matching) technique is used to adaptively update the probability $p_a(t)$ of each strategy a based on its known performance (frequently updated by the rewards received). Denote $ra(t)$ as the reward that a strategy a receives after its application at time t . $qa(t)$ is the known quality (or empirical estimate) of a strategy a , that is updated as follows

$$qa(t+1) = qa(t) + \alpha[ra(t) - qa(t)]$$

Where $\alpha \in (0, 1]$ is the adaptation rate. Based on this quality estimate, the PM method updates the probability $p_a(t)$ of applying each operator as follows

$$p_a(t+1) = p_{min} + (1 - K \cdot p_{min}) \frac{qa(t+1)}{\sum_{i=1}^K qa(t+1)}$$

Where $p_{min} \in (0, 1)$ is the minimal probability value of each strategy, used to ensure that no operator gets lost.

1) "DE/rand/1":

$$v_i = x_{r_1} + F \cdot (x_{r_2} - x_{r_3}) \quad (10)$$

2) "DE/rand/2":

$$v_i = x_{r_1} + F \cdot (x_{r_2} - x_{r_3}) + F \cdot (x_{r_4} - x_{r_5}) \quad (11)$$

3) "DE/rand-to-best/2":

$$v_i = x_{r_1} + F \cdot (x_{best} - x_{r_1}) + F \cdot (x_{r_2} - x_{r_3}) + F \cdot (x_{r_4} - x_{r_5}) \quad (12)$$

4) "DE/current-to-rand/1":

$$v_i = x_i + F \cdot (x_{r_1} - x_i) + F \cdot (x_{r_2} - x_{r_3}) \quad (13)$$

B. Combining Road and Ring topology

Topology information and mobility prediction are main distinction between the proposed algorithms and is exploitation of road topology map of the destination hotspot within a cellular network. Road network topology is represented by a stochastic finite state machine that is adopted directly from the digital map. Every edge in the map is represented by one state ki for each driving direction. Road topology is reorganization and management of node parameters and modes of operation from time to time to modify the network with the goal of extending its Lifetime while preserving other important characteristics. A ring network is a network topology in which each node connects to exactly two other nodes, forming a single continuous pathway for signals through each node - a ring.

C. Prefecting techniques:

Mobility prediction is an important maneuver that determines the location of the mobile terminal by carefully manipulating the available information. The prediction

accuracy depends on the user movement model and the prediction algorithm used. Two different techniques are already available for mobility prediction. First technique uses the historical movement patterns of the user to predict the user's whereabouts in future. The second technique uses the contextual knowledge. So, An adaptive algorithm is designed to predict the future positions of mobile terminals in terms of location probabilities based on moving directions and residence time in a cell. It makes use of LAR and DREAM routing technique to find the future mobile states of positions respectively.

D. Proposed Architecture: Algorithm



Figure 1. Activity mining with metastrategies

For performance evaluation, we have developed a suite of simulation tools using Java language on WinTel platform. The layer architecture of the simulator is depicted in Fig. 1. The simulation engine is a set of routines to carry out the operations designated by other modules. The mobile environment simulator is responsible for generating the network topology, supported services, and user behaviours. The users get their behaviour models from a behaviour pool. Each user has a number of behaviours to choose from at random. Each behaviour is associated with a lifetime. When the lifetime of all behaviours expires, a new set of behaviours is requested from the pool. To keep the behaviours fresh, each model has a lifetime in pool to trigger the generation of new models. Behaviours are selected from the pool following Zipf distribution.

E. Implementation of ZRP and AZRP

The **Adaptive Zone Routing Protocol**, as its name implies, is based on the concept of zones. A routing zone is defined for each node separately, and the zones of neighboring nodes overlap. The routing zone has a radius expressed in hops. The zone thus includes the nodes, whose distance from the node in question is at most n hops, where the routing zone of S includes the nodes A–I, but not K. In the illustrations, the radius is marked as a circle around the node in question. It should however be noted that the zone is defined inhops,

not as a physical distance. The nodes of a zone are divided into peripheral nodes and interior nodes. Peripheral nodes are nodes whose minimum distance to the central node is exactly equal to the zone radius. The nodes whose minimum distance is less than n are interior nodes. The nodes A–F are interior nodes; the nodes G–J are peripheral nodes and the node K is outside the routing zone. Note that node H can be reached by two paths, one with length 2 and one with length 3 hops. The node is however within the zone, since the shortest path is less than or equal to the zone radius. The number of nodes in the routing zone can be regulated by adjusting the transmission power of the nodes. Lowering the power reduces the number of nodes within direct reach and vice versa. The number of neighboring nodes should be sufficient to provide adequate reach ability and redundancy. On the other hand, a too large coverage results in many zone members and the update traffic becomes excessive. Further, large transmission coverage adds to the probability of local contention. AZRP refers to the locally proactive routing component as the **Adaptive Intra-zone Routing Protocol (AIARP)**. The globally reactive routing component is named **Adaptive Inter-zone Routing Protocol (AIERP)**. AIERP and AIARP are not specific routing protocols. Instead, AIARP is a family of limited-depth, proactive link-state routing protocols. AIARP maintains routing information for nodes that are within the routing zone of the node. Correspondingly, AIERP is a family of reactive routing protocols that offer enhanced route discovery and route maintenance services based on local connectivity monitored by AIARP. The fact that the topology of the local zone of each node is known can be used to reduce traffic when global route discovery is needed. Instead of broadcasting packets, AZRP uses a concept called *border casting*. Border casting utilizes the topology information provided by AIARP to direct query request to the border of the zone. The border cast packet delivery service is provided by the Border cast Resolution Protocol (BRP). BRP uses a map of an extended routing zone to construct border cast trees for the query packets. Alternatively, it uses source routing based on the normal routing zone. By employing *query control* mechanisms, route requests can be directed away from areas of the network that already have been covered.

V. CONCLUSION AND FUTURE WORK

We proposed Personalized activity mining and data management are expected to provide Even higher quality services than the proposed schemes. Obtaining rich Resource station and better equipped client services The cold start problem owing to the need for an initial behavior log DB remains unsolved. We work on that by dynamic tree structure and association mining. The Activity Prioritization methods can be improved and enables improved service Quality. Different Management strategies used for exploring activities in mobile computing. AZRP can be regarded as a routing framework rather than as an independent protocol. AZRP reduces the traffic amount compared to pure proactive or reactive routing. Routes to nodes within the zone are immediately available. AZRP makes an extension for ZRP

protocol that can adapt well to the complicated network with nodes moving non-uniformly. AZRP utilizes the excellent performance of the hybrid-driven manner of ZRP and simultaneously overcomes the bad adaptability of ZRP which assumes each node move uniformly and presets the same zone radius. For the mobility of nodes is variable in the practical networks, our future work may focus on the change of the zone radius aroused by the mobility change of nodes. This will be more accordant with the reality.

REFERENCES

- [1] G. Resta and P. Santi, "WiQoS: An Integrated QoS-Aware Mobility and User Behavior Model for Wireless Data Networks," *IEEE Trans. Mobile Computing*, vol. 7, no. 2, pp. 187- 198, Feb. 2008.
- [2] A. Yamasaki, H. Yamaguchi, S. Kusumoto, and T. Higashino, "Mobility-Aware Data Management on Mobile Wireless Networks," *Proc. IEEE 65th Vehicular Technology Conf.*, pp. 679-683, 2007.
- [3] J.-L. Huang and M.-S. Chen, "On the Effect of Group Mobility to Data Replication in Ad-Hoc Networks," *IEEE Trans. Mobile Computing*, vol. 5, no. 5, pp. 492-507, May 2006.
- [4] R.V. Mathivaruni and V. Vaidehi, "An Activity Based Mobility Prediction Strategy Using Markov Modeling for Wireless Networks" *Proceedings of the World Congress on Engineering and Computer Science 2008 WCECS 2008*, October 22 - 24, 2008, San Francisco, USA
- [5] I.F. Akyildiz and W. Wang, "The Predictive User Mobility Profile Framework for Wireless Multimedia Networks," *IEEE/ACM Trans. Networking*, vol. 12, no. 6, pp. 1021-1035, Dec. 2004.
- [6] I.F. Akyildiz and W. Wang, "The Predictive User Mobility Profile Framework for Wireless Multimedia Networks," *IEEE/ACM Trans. Networking*, vol. 12, no. 6, pp. 1021-1035, Dec. 2004.
- [7] W.-S. Soh and H. Kim, "QoS Provisioning in Cellular Networks Based on Mobility Prediction Techniques," *IEEE Comm. Magazine*, vol. 41, no. 1, pp. 86-92, Jan. 2003.
- [8] W. Su, S.-J. Lee, and M. Gerla, "Mobility Prediction and Routing in Ad Hoc Wireless Networks," *Int'l J. Network Management*, vol. 11, no. 1, pp. 3-30, Jan. 2001.